



ILLINOIS

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

PRODUCTION NOTE

University of Illinois at
Urbana-Champaign Library
Large-scale Digitization Project, 2007.

370.152
T2261
No. 456

STX

Technical Report No. 456

**WORD SCHEMAS: WHAT DO PEOPLE KNOW
ABOUT WORDS THEY DON'T KNOW?**

**William E. Nagy
and
Judith A. Scott**

University of Illinois at Urbana-Champaign

February 1989

Center for the Study of Reading

TECHNICAL REPORTS

The Library of the

APR 25 1989

University of Illinois
of Urbana-Champaign

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

174 Children's Research Center

51 Gerty Drive

Champaign, Illinois 61820

CENTER FOR THE STUDY OF READING

Technical Report No. 456

WORD SCHEMAS: WHAT DO PEOPLE KNOW ABOUT WORDS THEY DON'T KNOW?

**William E. Nagy
and
Judith A. Scott**

University of Illinois at Urbana-Champaign

February 1989

**University of Illinois at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820**

The work upon which this publication was based was supported in part by the Office of Educational Research and Improvement under Cooperative Agreement No. OEG 0087-C1001, with the Reading Research and Education Center. The publication does not necessarily reflect the views of the agency supporting the research.

EDITORIAL ADVISORY BOARD
1988-89

Beck, Diana

Commeyras, Michelle

Foertsch, Daniel

Hartman, Doug

Jacobson, Michael

Jehng, Jihn-Chang

Jimenez, Robert

Kerr, Bonnie

Kerr, Paul

Meyer, Jennifer

Moran, Juan

Ohtsuka, Keisuke

Roe, Mary

Schommer, Marlene

Scott, Judy

Stallman, Anne

Wilkinson, Ian

Wolff, Phillip

MANAGING EDITOR
Mary A. Foertsch

MANUSCRIPT PRODUCTION ASSISTANTS
Delores Plowman
Nancy Diedrich

Abstract

This study explores the acquisition of word schemas, that is, implicit knowledge about patterns of word meaning in English which constrain the hypotheses that proficient word learners make about new words they encounter. Subjects in 7th and 10th grade, and college undergraduates, were asked to rate the plausibility of made-up definitions for nonsense words, some of which conformed to constraints on the semantic composition of English words, and others of which violated these constraints in various ways. Results show that by the time they reach college, students have gained an implicit knowledge of word schemas reflecting subtle regularities of English vocabulary, which they can use when evaluating possible meanings for new words.

WORD SCHEMAS: WHAT DO PEOPLE KNOW ABOUT WORDS THEY DON'T KNOW?

Children learn new words at a phenomenal rate. According to recent published estimates, the average child may be learning as many as 3,000 (Nagy, Anderson, & Herman, 1987) or even 5,000 (Miller & Gildea, 1987) words a year during the school years. The rate at which children learn new words is even more astonishing when one considers that most of this learning must take place incidentally (Jenkins & Dixon, 1983; Nagy & Anderson, 1984). Children must infer the meanings of words from available oral or written contexts, which generally supply only limited information about the meaning of the word (Schatz & Baldwin, 1986).

Word learning, like other aspects of language acquisition, involves making and testing hypotheses on the basis of the available evidence. The question of how children learn words so quickly is therefore a specific case of the more general problem of induction in language acquisition, that is, how children manage to learn their language on the basis of what appears to be insufficient data.

Quine (1960) has provided a definitive example of the problem of induction as it relates to word learning.¹ In the situation Quine asks us to consider, the linguist investigating a new language hears a speaker of that language utter the word *gavagai* as a rabbit runs by. The problem facing the linguist is to constrain the hypotheses that can be made about the meaning of *gavagai*. This word might mean "rabbit," or it might mean "run," or even "running rabbit." One might think that a few simple tests could determine which of these was the actual meaning of *gavagai*. But the situation is not simple. There are an indefinitely large number of hypothetical word meanings that would fit the situation--for example, "rabbit ear," "triangle formed by two humans and a rabbit," "quadruped in the morning," "mammal moving northward," or "fur within five meters."

One cannot solve this problem simply by waiting to hear the word in another context, to see which of these hypotheses still fit. As Quine points out, there are indefinitely many possible meanings compatible with *any* context in which the word might be used, for example, "undetached rabbit part." Furthermore, children have been shown to acquire substantial, even if incomplete, information about the meanings of words on the basis of a single encounter in context (Carey, 1978; Dickinson, 1984; Nagy, Anderson & Herman, 1987; Nagy, Herman, & Anderson, 1985). Given the speed and efficiency with which children learn words, it is clear that there must be some restrictions on the hypotheses they make about the meanings of new words.

One might object that some of the hypotheses about *gavagai* suggested--"undetached rabbit part," "quadruped in the morning," or "mammal moving northward"--are ridiculous. That is exactly the point. The question is, do children learning their language know what constitutes a sensible hypothesis about the meaning of a new word, and if so, where do they get this knowledge?

Some psycholinguists (Keil, 1979, 1981; Markman, 1984; Markman & Hutchinson, 1984) have attempted to demonstrate the existence of constraints on the hypotheses young children make about word meanings. One such constraint is the assumption of taxonomic organization, a constraint investigated by Markman and Hutchinson (1984). In a series of experiments, Markman and Hutchinson showed that although young children favor thematic relationships in sorting and other tasks, they focus on taxonomic relationships when making hypotheses about the meanings of unfamiliar object names. That is, in most tasks, a child would group "cow" with "milk" rather than with "pig." But when given a nonce word as a new label for a picture of a cow, and asked what else could be named by that nonce word, children preferred a picture of a pig over one depicting milk. This presumably reflects the fact that there are many object names for taxonomically organized classes, but there are few, if any, words whose meanings incorporate thematically related entities such as "cow and milk" or "spider and web."

The case for some sort of limitations on the hypotheses children make about the meanings of new words is compelling; however, the nature and source of these limitations is not at all clear. Carey (1983), although agreeing that some constraints on word meanings are necessary to account for the ease and speed with which children learn words, presents strong arguments against the specific type of constraints suggested by Keil (1979), and suggests that no purely formal semantic constraints on word meanings are likely to work. Nelson (in press) argues that the constraint investigated by Markman and Hutchinson (1984) cannot be properly labeled a "constraint," because Markman and Hutchinson's data show only that children have a bias against, not an absolute avoidance of, thematically-based word meanings. Nelson also argues against the claim that the assumption of taxonomic organization is innate.

In this paper, our interest is in the later stages of language acquisition, where hypotheses about the meanings of new words are restricted not by the type of general constraints discussed by Keil (1979, 1981), Markman (1984), and Markman and Hutchinson (1984), but by the learner's knowledge of how his or her language typically organizes word meanings.

The assumption of taxonomic organization and other constraints discussed by Markman (1984) appear to be operative at early stages of language learning, and may well be universal heuristics, if not constraints, applied by children learning all languages. It is unlikely that differences in the knowledge of such constraints contribute very much to differences in word learning proficiency in the school years. To account for the rapid vocabulary growth that occurs throughout the school years, and for the large individual differences that emerge (Graves & Slater, 1987; Templin, 1957), we hypothesize that children implicitly recognize, to varying degrees, redundancies and regularities in the lexicon of their language, and utilize their knowledge of these regularities in learning new words. In order to avoid the connotations of innateness and universality associated with the word "constraints," we will borrow and build on the term "schema" from cognitive psychologists, and describe knowledge of systematicity in the lexicon in terms of "word schemas."

Research in cognitive psychology during the last 20 years has made it increasingly clear how important prior knowledge is in understanding and learning new information (see Anderson & Pearson, 1984, for a review). How people learn and understand new information depends crucially on their existing knowledge structures. However, although a strong effect of conceptual or content knowledge on word learning has been found (Nagy, Anderson & Herman, 1987), there has been no discussion of the role of "word schemas." "Word schemas" are analogous to "story structures." We hypothesize that just as knowledge about the nature and typical organization of stories aids the comprehension of new stories (Stein & Trabasso, 1982), knowledge about the nature of words and the organization of the lexicon aids the understanding of new words.

What kind of things do people know about words that might contribute to their ability to learn words? Clearly, good readers know something about the morphology of English--about prefixes, roots, and suffixes. But proficient word learners also have a great deal of knowledge about patterns of word meaning in their language. Along with knowledge of individual words, children acquire knowledge of basic semantic relationships such as synonymy, antonymy, hyponymy, and the part-whole relation (Kuczaj, 1982). They come to recognize recurrent patterns of polysemy, such as the causative and noncausative senses of words like *open*, *close*, *break*, and *run* (Bowerman, 1982). They also recognize patterns of meaning in groups of related words--semantic domains, or semantic sets (Kuczaj, 1982; Lyons, 1977). For example, speakers of English know, at least implicitly, that English verbs of motion typically tell something about the way an object moves (e.g., *slide*, *wobble*, *plunge*, *spin*), but not the shape of the object that is moving. This is in contrast to some languages which might have one verb meaning "to fall" which applies only to long rigid objects, and another totally different verb "to fall" that is reserved for round objects (see Friedrich, 1970, for a discussion of the important role that shape plays in some grammatical systems). Other languages may convey manner of motion through adverbs, as can English, but not through simple verbs such as *slide*, *stagger* and *slink*. Each semantic domain--verbs of motion, verbs of speaking, verbs of thinking, and so on--has characteristic ways in which

meanings are organized; the pattern of organization for a given semantic domain may differ substantially from one language to another.

Our fundamental hypothesis, then, is that word schemas (that is, knowledge about words, and especially, knowledge of how semantic information is organized into word meanings) play an important, and in fact indispensable, role in skilled word learning. Since the ultimate motivation for this research is to account for some of the individual differences that exist in word learning ability, we are interested particularly in knowledge about words children may gain as they are exposed to an increasing number of words, which may result in a kind of bootstrapping, or "Matthew effect" (Stanovich, 1986) in word learning.

The specific purpose of this study is to document the acquisition of word schemas reflecting implicit knowledge of fairly subtle regularities in the English lexicon--regularities that can only be recognized if one has a fairly substantial vocabulary. We know of only one other study, Nagy and Gentner (1987), that has looked at such word schemas. Nagy and Gentner included some constraints that were likely to be based on a fairly subtle knowledge of English vocabulary and found that adults applied implicit knowledge of constraints on possible word meanings in making and evaluating hypotheses about unfamiliar words encountered in context. However, only a few such constraints, only a few items per constraint and only adults were examined, so a replication and extension of this work was deemed necessary. In addition, it is important to explore the same type of knowledge with a different type of experimental measure, to provide convergent evidence concerning implicit knowledge of patterns of word meaning.

The task adopted in this experiment was inspired by the game "Dictionary" (or "Fictionary"). In this game, one team chooses an unfamiliar word from the dictionary, and in addition to its real definition, makes up other definitions as foils. The other team's job is to guess which definition is the real one. This game appears to depend on exactly the type of knowledge at issue here--the players' implicit sense of what is a possible (or likely) meaning for a new word. In the version of the task used in this experiment, instead of choosing among definitions, subjects were asked to rate definitions on a scale of how likely it was that a given definition could be a real definition, from "not very likely" to "very likely." If subjects can evaluate the plausibility of definitions they must be depending on knowledge of how word meanings are structured in English.

Pilot work with structured interviews suggested that, not surprisingly, even highly educated adults are generally unable to articulate their word schemas. An important goal in the design of this experiment was therefore to construct definitions which violated only very specific hypothesized components of people's word schema knowledge, and were otherwise well-formed, possible word meanings in English.

Method

Subjects

Subjects were 42 seventh-grade students, 37 tenth-grade students and 42 college undergraduates. The undergraduates were students at a large midwestern university, fulfilling a course requirement. The 7th and 10th graders were students in a small rural community.

Materials

The task was to rate 96 definitions. Each item consisted of a word, its part of speech (noun or verb), a definition, and a sentence illustrating how that word would be used. The subjects were asked to rate, on a 4-point scale, the likelihood that the definition given represented the real definition for the word. Table 1 gives a sample item.

[Insert Table 1 about here.]

Most of the words and definitions were made up. Dictionaries of unusual words (Byrne, 1974; Dickson, 1984) served as a source for some low-frequency words, and sometimes as an inspiration for implausible meanings. Noun and verb versions of each item were constructed with essentially the same content, differing only in the part of speech listed, minor details in the definition, and the syntax of the example sentence.

Three categories of items were developed; there were 32 items in each of these categories. The first category, which constitutes the core of this experiment, we will refer to as "ill-formed." Items in this category were constructed so that their meanings were inconsistent with certain regularities we postulate to hold for English verbs. Thus, the verb version of items in this category did not constitute plausible meanings in English. However, the noun versions of these items did constitute possible English meanings.

This category contained several subtypes of items. One subtype contained "object incorporation" items. This subtype takes advantage of the fact that, in English, there are nouns that include the meaning of a verb plus its object, for example, *fratricide*, murder of one's brother, or *polytheism*, belief in many gods. But there are no verbs corresponding to such nouns; for example, there is no word meaning "to murder one's brother" or "to believe in many gods." For each item, two versions were constructed: a noun version, which was a possible meaning in English, and a verb version, which was a type of meaning that does not typically occur in English.

Other subtypes of the ill-formed category were similarly designed so that although the noun and verb versions of the item were almost identical in content, the verb version constituted an ill-formed or implausible meaning, whereas the noun version constituted a plausible meaning. For example, in the "passive" subtype of ill-formed items, one item had as a verb meaning "to be hit on the head," with the corresponding noun meaning "a blow to the head." The verb meaning, "to be hit on the head," is not a likely meaning in English; English verbs do not incorporate the passive voice in this manner.²

The differences between the noun and verb versions of items in the "ill-formed" category are the focus of this experiment. These differences reflect relatively subtle constraints on the patterning of word meanings in English. The noun and verb versions are identical in conceptual content, thus controlling for a variety of other factors (e.g., the utility of having a word for the concept of watching the sunrise or hitting someone on the head).

The second main category of items is labeled "well-formed." These items represent possible meanings in English, and should be equally plausible as nouns or verbs. These items serve primarily as a baseline, to deal with the possibility that nouns and verbs might receive differential ratings independent of their content.

The third main category of items consisted of items with meanings that were inherently implausible in terms of their conceptual content (e.g., "to meditate underwater"), regardless of their part of speech. For the purposes of this study, these items serve simply as fillers.

Table 2 gives examples of the three basic categories of items, along with specific subtypes within these categories.

[Insert Table 2 about here.]

Another question raised by Nagy and Gentner (1987) concerned the effect of morphology on constraints. In that study, Nagy and Gentner looked at some constraints that involved the morphological structure of words, such as the fact that meanings involving proper names must have the proper name morphologically present in the word. (For example, the word *Frenchify* meaning "to make French, to translate into French," is more plausible than the word *endorate* with the same

meaning.) However, they did not systematically explore the relationship between constraints on word meanings and the morphological transparency of words. In this study, morphological transparency and well-formedness of meaning are treated as orthogonal variables, in order to see what interaction, if any, there is between the two. The hypothesis is that constraints on word meanings will apply more strictly to words when the meaning is not conveyed by the morphological structure of the word.

Part of speech and morphological complexity were counter-balanced. Thus, for each item there were four versions: a morphologically complex noun, a morphologically complex verb, a morphologically simple noun and a morphologically simple verb. There were four versions of the test, so that a subject only saw one version of each item. Table 3 gives the four versions of an ill-formed item, in which the noun version is more acceptable as a word meaning than the verb version.

[Insert Table 3 about here.]

After the body of the rating task, subjects were given open-ended questions concerning the strategies they had utilized in rating the items, and then a task in which they rated their reliance on specific types of information (see Table 4).

[Insert Table 4 about here.]

Procedures

The 7th and 10th grade students were tested in their classrooms. The undergraduates were tested in small groups. All of the subjects completed a timed, wide-range multiple-choice vocabulary test prior to the definition task. The 7th-grade vocabulary test was scaled down and only incorporated two pages of the test taken by the 10th graders and undergraduates. After the vocabulary test, the subjects completed the definition rating task and self-reported strategy questions at their own pace.

Design & Analysis

The main analysis was a $3 \times 2 \times 2 \times 2$ mixed analysis of variance, with grade (7th, 10th or undergraduate) as a between-subject factor and well-formedness (well-formed or ill-formed), morphological complexity (complex or simple), and part of speech (noun or verb) as within-subject factors. Post hoc analyses looked at effects within each grade.

Our primary interest is the difference between the noun and verb versions of ill-formed items. Because the well-formed items serve as a baseline for any possible overall difference in ratings between nouns and verbs, our prediction is more specifically that there will be a significant Well-Formedness \times Part of Speech interaction.

A secondary analysis was performed on the undergraduate data using a hierarchical multiple regression procedure following the logic of mixed analysis of variance. This analysis examined the contribution of vocabulary (as measured by the wide range vocabulary test) and strategies (as measured by the strategy rating scale). The dependent measure of each analysis was the rating of likelihood that the item represented a real definition.

Results and Discussion

The results of the ANOVA indicate significant main effects for morphological complexity $F(1,120) = 72.84, p < .001$; well-formedness $F(1,120) = 109.5, p < .001$; and part of speech $F(1,120) = 16.10, p < .001$. Means and standard deviations for each cell are presented in Table 5.

[Insert Table 5 about here.]

There was no main effect for grade, $F(2,120) = 1.56, p > .05$, which indicates that subjects in different grades were equivalent in their overall use of the rating scale.

The strong main effect of morphological complexity shows that subjects rated the complex, semi-transparent words, (for example, *motocept*) higher than the simple, non-transparent words (for example, *lisk*). A significant Grade x Morphological Complexity interaction, represented in Figure 1, $F(2,120) = 17.23, p < .001$, indicates that subjects at different grade levels were not equivalent in their use of morphological complexity to determine the acceptability of definitions. Post hoc Tukey tests with $\alpha = .05$ indicate that 7th grade students did not differ in their ratings for morphologically simple and morphologically complex words. However, the difference in mean ratings between simple words and complex words is significant both for 10th-grade and undergraduate subjects.

[Insert Figure 1 about here.]

A significant main effect of well-formedness, $F(1,120) = 109.49, p < .001$, indicates that students rated ill-formed items as less likely than well-formed items. The main effect of part of speech was also significant, $F(1,120) = 16.00, p < .001$. However, these main effects must be interpreted in light of the significant interaction of well-formedness with part of speech, $F(1,120) = 19.45, p < .001$.

[Insert Figure 2 about here.]

This predicted interaction between well-formedness and part of speech, seen in Figure 2, is the chief finding. This significant interaction indicates that subjects rated ill-formed verb items lower than ill-formed noun items, apart from any overall rating differences that might exist between nouns and verbs. Thus, subjects were sensitive to the subtle regularities in the pattern of English word meanings that distinguished the verbs and nouns.

Remember that the noun and verb versions of each item are virtually identical in conceptual content, and differ only in the surface syntactic form. So, conceptual plausibility is not at issue here. Thus, the Well-Formedness x Part of Speech interaction represents specifically linguistic knowledge, and in fact, very subtle knowledge about how English part-of-speech categories differ in the type of information they can include.

Morphological complexity interacted with only one within-subject variable, well-formedness $F(1,120) = 21.88, p < .001$, suggesting that constraints on word meanings apply more strictly to words which are not perceived as morphologically transparent. Figure 3 depicts the morphological complexity by well-formedness interaction. The difference between the well-formed and ill-formed items was greater for the morphologically simple versions than for the morphologically complex versions. None of the three-way interactions involving morphological complexity were significant.

[Insert Figure 3 about here.]

There were interactions between grade and part of speech, $F(2,120) = 4.90, p < .01$, and between grade and well-formedness $F(2,120) = 5.89, p < .01$, and also a three-way interaction between grade, well-formedness, and part of speech, $F(2,120) = 4.19, p < .05$. These interactions are depicted in Figure 4. Essentially, these interactions indicate that the Well-Formedness x Part of Speech interaction discussed above becomes more pronounced with the older subjects. Post hoc Tukey tests, with $\alpha = .05$, reveal that only the undergraduates differed significantly in their rating of nouns ($M = 2.38$) and verbs ($M = 2.01$) for ill-formed items. Thus, the sensitivity to subtle regularities evidenced by differences in rating between nouns and verbs for ill-formed items was not reliably present at 7th or 10th grade.

[Insert Figure 4 about here.]

However, further examination of the Grade x Well-Formedness interaction, using Tukey tests with $\alpha = .05$, found that there were significantly different ratings for the ill-formed and well-formed items in all three grades. These results, shown in Figure 5, indicate that even though the 7th graders made less of a distinction in rating ill-formed and well-formed items than the 10th graders or the undergraduates, they still have sensitivity to some factors that contribute to the plausibility of word meanings.

[Insert Figure 5 about here.]

Subjects' Self-Reported Strategies

Table 6 gives the intercorrelations between the self-reported strategies students used to rate the definitions (see Table 4). The strategy of thinking of similar words is correlated with attention to word parts and thinking about the utility of the word. Whether the definition sounded realistic was correlated with attention to the utility of the word and the naturalness of the sentence. Attention to word parts was not correlated with any of the other strategies.

[Insert Table 6 about here.]

A 3 x 6 mixed analysis of variance was performed with grade (7th, 10th and college) as a between-subject variable and strategy type (the questions in Table 4) as a within-subject variable. Table 7 gives the means and standard deviations of subject responses for the six self-reported strategies. There was a significant effect of strategy type $F(5,114) = 9.90, p < .001$, but neither the main effect of grade nor the interaction of grade with strategy type were significant ($p > .05$). This analysis indicates that subjects reported relying on some types of information more than others in rating the definitions. For example, they reported relying relatively little on the possible utility of the word or the way the word sounded. More attention was paid to the meanings of the parts of the words and how natural the sentence sounded. However, the relative ratings for these different questions is not the primary interest. Rather, we are interested in how variation among students in their answer to individual strategies relates to their performance on rating the definitions. For example, one would expect that students who said they paid attention to word parts would make a greater differentiation between the morphologically complex and morphologically simple items.

[Insert Table 7 about here.]

Additional Analyses on the Undergraduate Data

The type of strategies used to evaluate word meanings (as measured by the self-reported strategies at the end of the test), and verbal ability (as measured by the vocabulary test), were considered as factors which might interact with word schema knowledge. However, the results show that some of the subtle linguistic knowledge we are measuring is late in developing and only reliably present in the undergraduate sample. We therefore limited these analyses to the data from the undergraduate subjects. The data was analyzed using a hierarchical multiple regression procedure following the logic of mixed analysis of variance.

[Insert Table 8 about here.]

The results given in Table 8 indicate that the main effects of well-formedness, morphological complexity, and part of speech, and interactions between well-formedness and syntax and well-formedness and morphological complexity parallel the results of the main analyses. The main effect of vocabulary indicates that people with bigger vocabularies rated the definitions as less likely to be meanings for real words than those with smaller vocabularies. However, those who paid attention to the naturalness of the definition sentence rated the items more highly overall.

Of special interest in this analysis are the interactions of the between-subject variables with the vocabulary test and the self-reported strategies. The significant Morphological Complexity x Vocabulary interaction indicates that undergraduates with bigger vocabularies made a greater distinction in ratings between morphologically simple and morphologically complex items.

There were also two interactions of morphological complexity with strategy debriefing questions. People who said they paid more attention to word parts did in fact make a greater distinction in ratings between morphologically simple and morphologically complex items. Also, people who said they paid attention to the possible utility of the word similarly made a greater distinction between morphologically simple and morphologically complex items.

Thus, the extent to which subjects took word parts into account as they rated the definitions seems to depend partly on their vocabulary knowledge, and partly on the attention that they paid to word parts or the utility of a word. Interestingly, vocabulary and attention to word parts had independent effects on the ratings. These two variables are not significantly correlated ($r = .08$). The order of entry into the equation did not make a difference. The Morphological Complexity x Vocabulary interaction is significant when entered after the Morphological Complexity x Attention to Word Parts interaction. When these were entered in the opposite order, both were still highly significant.

One surprising result was that attention to word parts also interacted with knowledge of the noun and verb differences for the ill-formed items (the Attention to Word Parts x Well-Formedness x Part of Speech interaction). This is especially interesting since the parallel interaction involving vocabulary knowledge instead of attention to word parts (Vocabulary x Well-Formedness x Part of Speech) did not even approach significance ($F < 1.0$). One interpretation for this result is that attention to word parts is symptomatic of a more general sensitivity to words. People who pay attention to word parts may be more analytic in their approach to word meanings and therefore more sensitive in general to recurrent patterns in the internal semantic structure of words. On the other hand, possession of a large vocabulary alone does not guarantee this type of sensitivity to words.

With the exception of the Attention to Utility x Morphological Complexity interaction, there are no other significant interactions between the self-reported strategies and the within-subjects variables (morphological complexity, part of speech, and well-formedness). Although the subjects said that they used other types of knowledge about words to rate them—for example, they considered whether the definition sounded realistic—this type of knowledge did not influence the ratings of any of the variables. This data and interviews with a subset of the subjects confirm our belief that only some aspects of word schema knowledge are normally accessible to conscious attention. Only a few of the subjects gave any indication that they were aware of the role of part of speech in their ratings, yet they obviously used this information to rate the definitions. On the other hand, many subjects indicated awareness of the role of word parts in their judgments.

One limitation of this study is that it does not directly address the role of word schemas in word learning. In fact, although there is clear evidence that older students recognize subtle regularities in the English lexicon, the results do not give any direct support for the hypothesis that such knowledge plays an important role in word learning. For example, we would have expected a significant correlation between attention to word parts and vocabulary size; those students who are more aware of word parts should be more efficient word learners, and hence have larger vocabularies. The correlation, though positive, was nonsignificant. Similarly, we would have predicted a significant Vocabulary x Well-Formedness x Part of Speech interaction, on the assumption that subjects who were more sensitive to the subtle regularities of English vocabulary being measured would in general be better word learners. However, this interaction did not even approach significance.

The most pessimistic interpretation of these results (from our perspective) would be that although the older subjects do show sensitivity to regularities in the lexicon, this sensitivity has nothing to do with vocabulary size, and hence, nothing to do with word learning ability. But this pessimistic conclusion,

although consistent with the results, is not the only permissible interpretation. There are at least two ways one can maintain the initial hypothesis, that word schema knowledge plays an important role in vocabulary growth, in light of the results.

First of all, a distinction must be made between vocabulary size and word learning ability. Although word learning ability contributes to vocabulary growth, we would expect *opportunity* for vocabulary growth--in terms of the volume of exposure to vocabulary-rich language--to make a far bigger contribution. Ability is not the only factor that determines success.

Second, a distinction must be made between tacit knowledge of regularities in the lexicon, and display of that knowledge on tasks such as the one used here. Subjects who made a distinction in ratings between nouns and verbs for the ill-formed items obviously possess tacit knowledge of the constraints on verb meanings. However, failure to make such a distinction in this experimental task does not in itself prove that the subject does not possess such tacit knowledge. Thus, the two-way interaction of Attention to Word Parts x Morphological Complexity, and the three-way interaction of Attention to Word Parts x Well-Formedness x Part of Speech, could be interpreted as indicating that subjects who are more aware of the morphological structure of words are more likely to display their tacit knowledge of other regularities in the lexicon on this rating task. Likewise, the interactions involving grade may reflect in part an increasing ability to deal with this kind of task, as well as increased knowledge of the structure of the lexicon.

However, the primary purpose of this experiment was not to investigate the relationship between word schema knowledge and word learning ability, but to demonstrate the existence of the tacit knowledge of a few subtle regularities in the English lexicon, and, by implication, to show that proficient word learners have substantial knowledge about patterns of word meaning in English which they can apply when dealing with new words.

Conclusion

This study has demonstrated that people do know a lot about words they do not know and that this knowledge increases from junior high school through college. At the college level, subjects' word schemas include recognition of very subtle differences in the permissible semantic content of nouns and verbs. The lexicon of English, despite its massive irregularities and arbitrary details, contains semantic patterns and regularities that skilled readers can recognize and apply in making judgments about the plausibility of new definitions.

The subtle syntactic knowledge measured in this study appeared to be tacit for most of the subjects and it was reliably present only for the undergraduate subjects. It is important to note that the subtle noun-verb differences found here must have been gained from exposure to relatively few and infrequent words in the language.³ It should also be noted that the word schemas investigated in this study presumably represent only a tiny fraction of the word schema knowledge people use when they are trying to figure out what a new word means.

Large scale vocabulary growth can only be achieved by increasing students' independent word learning (Nagy & Anderson, 1984, Sternberg, 1987). Most research on independent word learning has focused on characteristics of the text, such as context clues, or on broadly-defined measures of ability (see for example Sternberg, 1987; Sternberg & Powell, 1983). The research reported here suggests that word learning is dependent, not only on general ability, but also on specific knowledge about the lexical semantic structure of English.

References

- American Heritage School Dictionary* (1977). Boston, MA: Houghton Mifflin Co.
- Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of reading comprehension. In P. D. Pearson (Ed.), *Handbook of reading research* (pp. 255-291). New York: Longman.
- Bowerman, M. (1982). Reorganizational processes in lexical and syntactic development. In E. Wanner & L. R. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 319-346). Cambridge, MA: Cambridge University Press.
- Byrne, Mrs. (1974). *Mrs. Byrne's dictionary of unusual, obscure, and preposterous words*. Secaucus, NJ: University books, Inc.
- Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 264-293). Cambridge, MA: MIT Press.
- Carey, S. (1983). Constraints on word meaning. In T. Seiler & W. Wannenmacher (Eds.), *Concept development and the development of word meanings* (pp. 126-143). New York: Springer Verlag.
- Dickinson, D. (1984). First impressions: children's knowledge of words gained from a single exposure. *Applied Psycholinguistics*, 5, 359-373.
- Dickson, P. (1984). *Words: A connoisseur's collection of old and new, weird and wonderful, useful and outlandish words*. New York: Delacorte Press.
- Friedrich, P. (1970). Shape in grammar. *Language*, 46, 379-407.
- Graves, M. F., & Slater, W. H. (1987). *The development of reading vocabularies in children from three social, economic, and linguistic settings: a preliminary report*. Paper presented at the American Educational Research Association Annual Meeting. Washington, D.C.
- Jenkins, J., & Dixon, (1983). Vocabulary learning. *Contemporary Educational Psychology*, 8, 237-260.
- Kiel, F. (1979). *Semantic and conceptual development: An ontological perspective*. Cambridge, MA: Harvard University Press.
- Kiel, F. (1981). Constraints on knowledge and cognitive development. *Psychological Review*, 88, 197-227.
- Kuczaj, S. A. (1982). Acquisition of word meaning in the context of the development of the semantic system. In C. Brainerd & M. Pressley (Eds.), *Verbal processes in children: progress in cognitive development research* (95-123). New York: Springer-Verlag.
- Lyons, J. (1977). *Semantics*. Cambridge, MA: Cambridge University Press.
- Markman, E. (1984). *How children constrain the possible meanings of words*. Unpublished manuscript.
- Markman, E., & Hutchinson, J. (1984). Children's sensitivity to constraints on word meanings: Taxonomic versus thematic relations. *Cognitive Science*, 16, 1-27.
- Miller, G., & Gildea, P. (1987). How children learn words. *Scientific American*, 257(3), 94-98.

- Nagy, W., & Anderson, R. (1984). How many words are there in printed school English? *Reading Research Quarterly*, 19, 304-330.
- Nagy, W., & Gentner, D. (1987). *Semantic constraints on lexical categories* (Tech. Rep. No. 413). Urbana: University of Illinois, Center for the Study of Reading.
- Nagy, W., Herman, P., & Anderson, R. (1985). Learning words from context. *Reading Research Quarterly*, 20, 233-253.
- Nelson, K. (in press). Constraints on word learning? *Cognitive Development*.
- Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press.
- Schatz, E., & Baldwin, R. S. (1986). Context clues are unreliable predictors of word meanings. *Reading Research Quarterly*, 21, 439-453.
- Stanovich, K. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21(4), 360-406.
- Stein, N. L., & Trabasso, T. (1982). What's in a story: An approach to comprehension and instruction. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 2) (pp. 213-254). Hillsdale, NJ: Erlbaum.
- Sternberg, R. J. (1987). Most vocabulary is learned from context. In M. McKeown & M. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 89-105). Hillsdale, NJ: Erlbaum.
- Sternberg, R. J., & Powell, J. S. (1983). Comprehending verbal comprehension. *American Psychologist*, 38, 878-893.
- Templin, M. C. (1957). *Certain language skills in children: Their development and interrelationships*. (The Institute of Child Welfare, Monograph Series No. 26). Minneapolis: University of Minnesota.

Footnotes

¹Quine was actually not directly interested in the problem of language acquisition. However, his example, as Carey (1983) points out, is a perfect illustration of the need for some sort of constraints on the hypotheses that children make about the meanings of new words.

²A possible exception to this generalization is one meaning of the word *fall*, described in the *American Heritage School Dictionary* (1977) as "to be wounded or killed." It is arguable, of course, whether this is the best characterization of this meaning of *fall*.

³Margaret Richek (personal communication, June, 1988) pointed out to us that not all of the verbs we had considered to be ill-formed necessarily were. There are a few verbs in English that would fall into the category of morphologically transparent object incorporation, such as *stargaze* or *birdwatch*. Examination of the results for individual item subtypes indicates that "general object-incorporating verbs" (see Table 2) are, in fact, as acceptable as the corresponding nouns if they are morphologically transparent. For morphologically simple object incorporation items, the predicted part of speech difference does occur. For all other individual subtypes of "ill-formed" items there is an obvious difference in ratings between the nouns and the verbs, for both morphologically complex and morphologically simple versions.

We take this wrinkle in the data as further confirmation of the conclusion that undergraduates, at least, are sensitive to very subtle regularities in the lexicon, even those that are exemplified by only a very small number of words in the language. The undergraduate sample must have drawn generalizations about these patterns of word meanings from very little data.

Table 1**Sample Item from the Definition Rating Task**

HOW LIKELY IS IT THAT THIS IS THE REAL MEANING?

belligression (noun): movement toward someone in a hostile or unfriendly way.

"Wayne watched the belligression of the terrorist toward the banker."

very unlikely 1 2 3 4 very likely

Table 2**Item Categories Used in the Definition Rating Task****I. ILL-FORMED:**

1. Object incorporation: nouns are allowed to represent a verb plus an object but verbs can't incorporate objects.

-phobias: phobias are nouns, not verbs.

botanophobiate (verb): to fear plants and flowers.

"The kitten seemed to botanophobiate."

-isms: beliefs are nouns, not verbs

eidolize (verb): to believe in ghosts.

"From that moment on, she eidolized."

General object incorporation constraint.

solispect (verb): to watch the sunrise.

"Carol often gets up early to solispect."

2. 'Voice' constraints on intransitive verbs: Intransitive verbs do not convey properties, emotions, or passive voice.

fogate (verb): to be very foggy.

"The airport fogated, stranding passengers for four days."

indole (verb): to be very sad.

"Karen indoled when Roy refused to be her friend."

incapulate (verb): to be hit on the head.

"Harry incapulated sixteen times during the attack."

Table 2 (Continued)**II. FILLER ITEMS:**

dismuzacate (verb): to illegally deactivate the muzak system in a public place.

"They searched the walls for a way to dismuzacate the elevator."

III. WELL-FORMED:

1. No constraint, well-formed as verbs and nouns.

musticate (verb): to become moldy or musty.

"Her shoes began to musticate in the closet."

2. No constraint, well-formed as verbs and nouns which indicate motion plus manner.

belligress (verb): to move toward someone in a hostile or unfriendly way.

"The terrorist belligressed toward the banker."

3. No constraint, well-formed as verbs and nouns which indicate non-motion with an adverb of degree or manner.

permaciate (verb): to waste away gradually.

"Illness can cause a person to permaciate."

Table 3**Four Versions of an Ill-formed Item**

Morphologically Complex Noun

motoception (noun): the chasing of cars
The old dog dreamed about motoception.

Morphologically Complex Verb

motocept (verb): to chase cars
The dog motocepts everytime he gets a chance.

Morphologically Simple Noun

lisk (noun): the chasing of cars
The old dog dreamed about lisk.

Morphologically Simple Verb

lisk (verb): to chase cars
The dog lisks everytime he gets a chance.

Table 4**Form for Student's Self-Report of Strategies Used when Rating the Definitions**

Now that you have finished your rating task, we would like you to try to describe what you thought about as you read these items.

Please take a few minutes to answer the questions below:

1. Can you give us any rules or guidelines that you found yourself using as you tried to rate different items?
(Please write clearly!)

a. I gave an item a low rating if _____

b. I gave an item a high rating if _____

2. Other comments:

- II. Please circle the answers that reflect how much you relied on different types of information to rate items in this study:

1. I asked myself if it sounded like the meaning or definition of other words I could think of.

very little 1 2 3 4 a whole lot

2. I paid attention to what parts of the word meant.

very little 1 2 3 4 a whole lot

3. I thought about whether anyone would ever need a word like this word.

very little 1 2 3 4 a whole lot

4. I paid attention to the way the word sounded.

very little 1 2 3 4 a whole lot

Table 4 (Continued)

5. I considered whether the sentence seemed natural.

very little 1 2 3 4 a whole lot

6. I asked myself if the definition sounded like one I would actually find in a dictionary.

very little 1 2 3 4 a whole lot

- III. Pick the ONE statement (numbered 1 to 6 above) which tells what was the MOST important to you when you were rating the items.

Please put the number of that statement here _____

Table 5**Means and Standard Deviations for Definition Ratings**

		Morphological Complexity			
Grade		Complex		Simple	
		Nouns	Verbs	Nouns	Verbs
Seventh:	Well-formed meanings	2.58 (.45)	2.61 (.47)	2.54 (.51)	2.59 (.58)
	Ill-formed meanings	2.55 (.65)	2.45 (.53)	2.388 (.55)	2.33 (.55)
Tenth:	Well-formed meanings	2.61 (.54)	2.63 (.66)	2.50 (.59)	2.43 (.55)
	Ill-formed meanings	2.47 (.50)	2.33 (.54)	2.13 (.58)	2.01 (.48)
College:	Well-formed meanings	2.79 (.44)	2.75 (.50)	2.29 (.60)	2.34 (.54)
	Ill-formed meanings	2.75 (.49)	2.33 (.52)	2.01 (.59)	1.68 (.53)

Note: 1 = very unlikely to be a real definition
 4 = very likely

Table 6**Correlation Matrix for the Self-Report Strategy Questions in All Grades**

	(2)	(3)	(4)	(5)	(6)
(1) Thought of similar words	.3080**	.2641*	.0827	-.0051	.1860
(2) Attention to word parts		.1350	.1705	.0868	.0127
(3) Attention to word utility			.1035	.0874	.4387**
(4) Attention to word sounds				.1325	.1376
(5) Attention to the sentence					.2781*
(6) Whether definition sounded realistic					1.0000

Note. N = 115

* = $p < .01$

** = $p < .001$ (1 tailed significance test)

Table 7**Means (and Standard Deviations) for Self-Reported Strategy Questions**

		7th	10th	College	Total
(1)	Thought of similar words	2.85 (1.08)	3.11 (.84)	3.05 (.98)	3.00 (.98)
(2)	Attention to word parts	3.18 (.96)	3.17 (1.07)	3.55 (.63)	3.31 (.91)
(3)	Attention to word utility	2.41 (1.17)	2.61 (1.19)	2.86 (.97)	2.63 (1.12)
(4)	Attention to word sounds	2.92 (1.05)	2.89 (.92)	3.07 (.96)	2.97 (.98)
(5)	Attention to the sentence	3.31 (.72)	3.14 (1.01)	3.29 (.80)	3.25 (.85)
(6)	Whether definition sounds realistic	3.45 (.79)	3.39 (.83)	3.05 (.98)	3.29 (.89)

Note. 1 = Used this type of information very little
 4 = Used this type of information a whole lot

Table 8**Analysis of Undergraduate Data, Aggregated by Subject**

Variable	Regression Coefficients	% of Variance	Simple <i>R</i>	<i>F</i>
Between subjects				
Vocabulary	-.01	24.4	-.49	12.9**
Attention to the sentence	.13	8.2	.35	4.7*
Attention to word parts	-.04	0.5	-.05	0.4
Attention to utility	-.00	0.0	.00	0.0
Constant/residual	2.53	66.9		
Within Subjects				
Morphological ^a complexity	-.35	20.5	.45	220.2**
Well-Formedness ^b	.15	11.2	.27	80.5**
Part of Speech ^c	.21	3.3	.15	22.9**
Well-Formedness x Part of Speech	.16	3.2	-.15	22.9**
Well-Formedness x Morphological Complexity	-.06	1.4	-.09	9.7**
Vocabulary x Morphological Complexity	.01	3.0	.47	21.89**

Table 8 (Continued)

Variable	Regression Coefficients	% of Variance	Simple <i>R</i>	<i>F</i>
Attention to Word Parts x Morphological Complexity	.14	2.8	.47	20.0**
Attention to Utility x Morphological Complexity	0.04	0.6	.41	4.0
Attention to Word Parts x Part of Speech	-.03	0.2	.14	1.1
Attention to Word Parts x Well-Formedness	.01	0.0	.27	0.0
Attention to Word Parts x Well-Formedness x Part of Speech	-.07	0.7	-.16	5.3*
Constant/residual	2.53	51.2		

* $p < .05$ ** $p < .01$ ^aCoded +1 complex, -1 simple.^bCoded +1 well-formed, -1 ill-formed.^cCoded +1 nouns, -1 verbs.

Figure Captions

Figure 1. Grade by morphological complexity interaction.

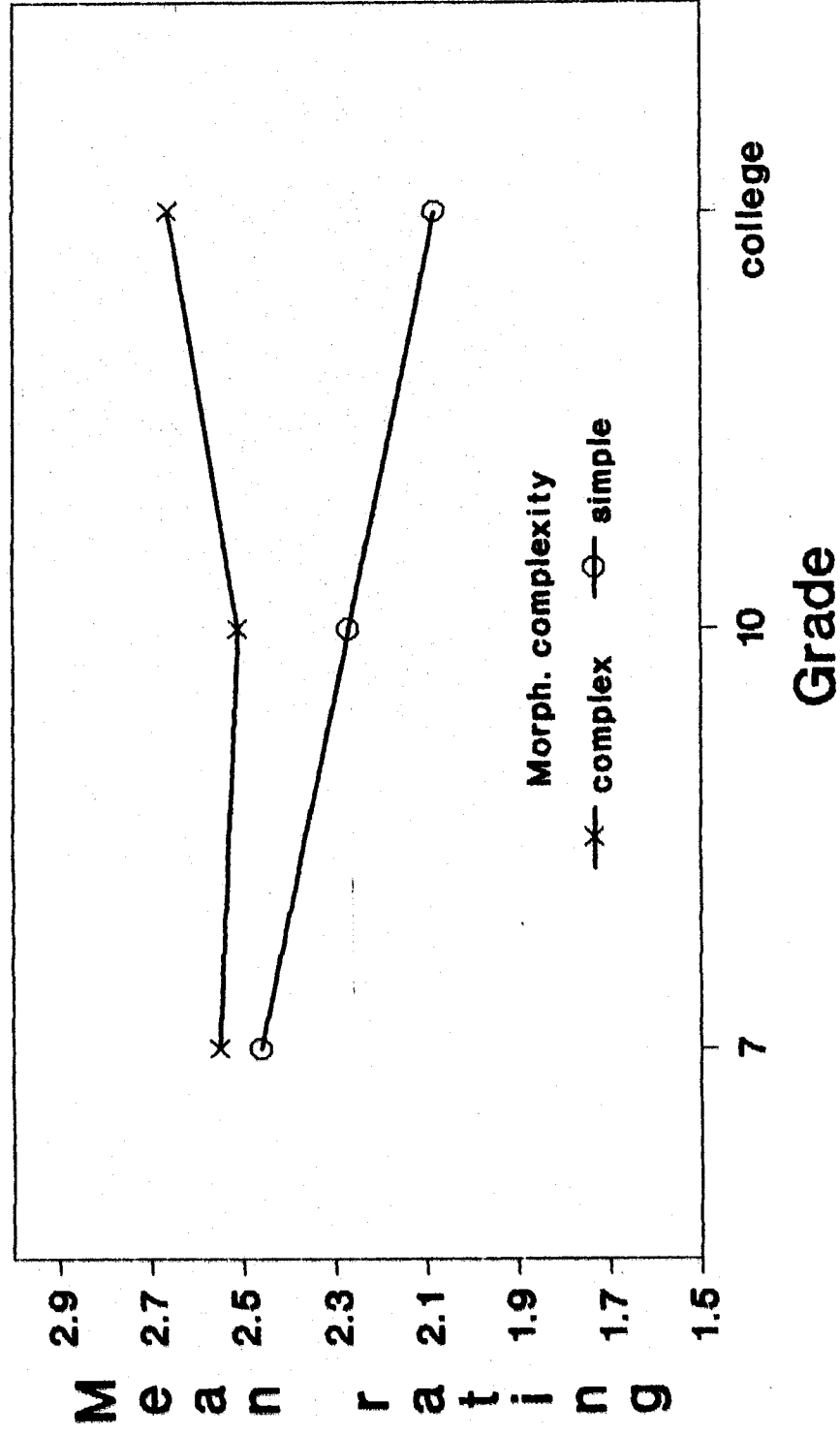
Figure 2. Well-formedness by part of speech interaction.

Figure 3. Morphological complexity by well-formedness interaction.

Figure 4. Grade by well-formedness by part of speech interaction.

Figure 5. Grade by well-formedness interaction.

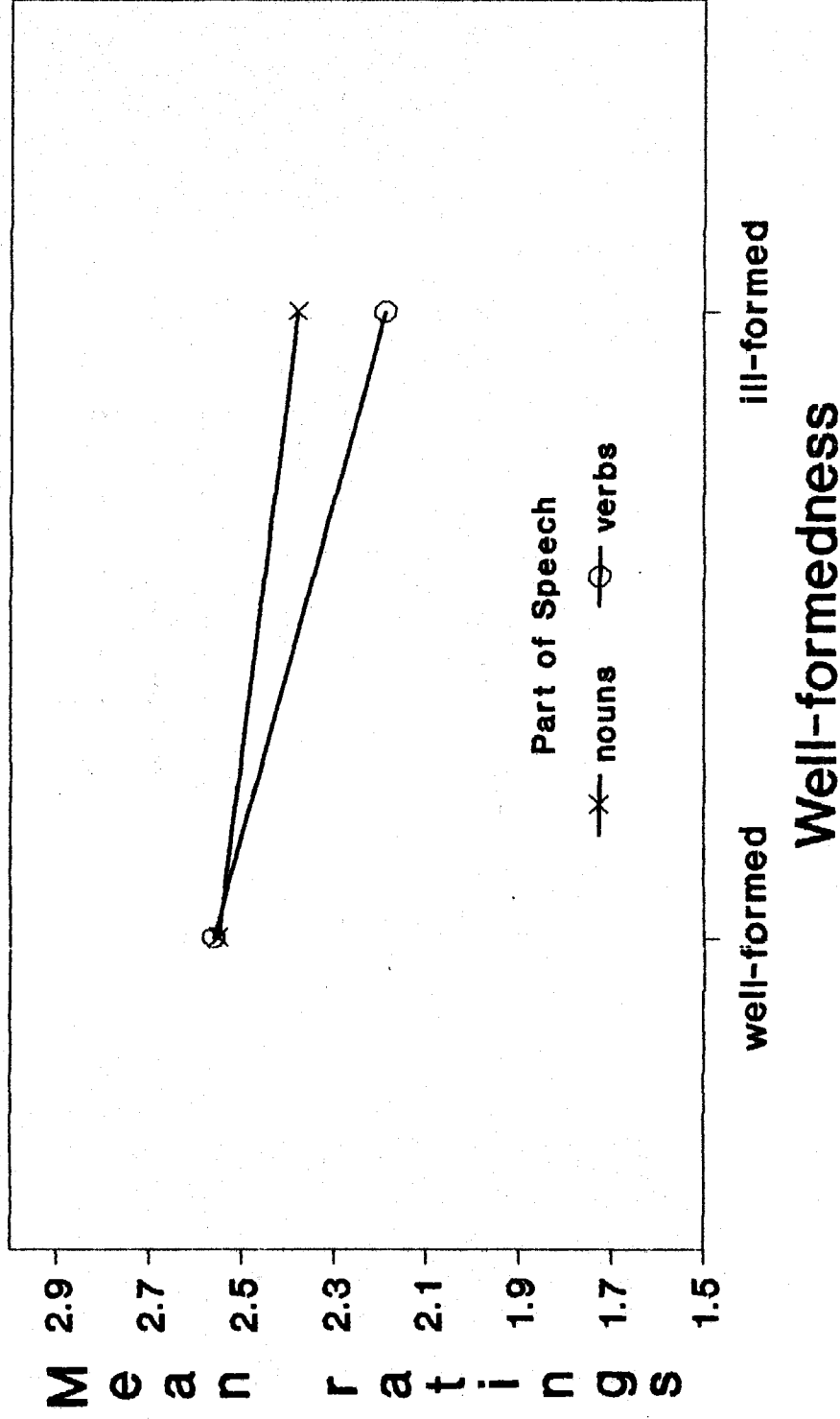
Grade by morphological complexity Interaction



Ratings: 4=very likely 1=very unlikely

Figure 1

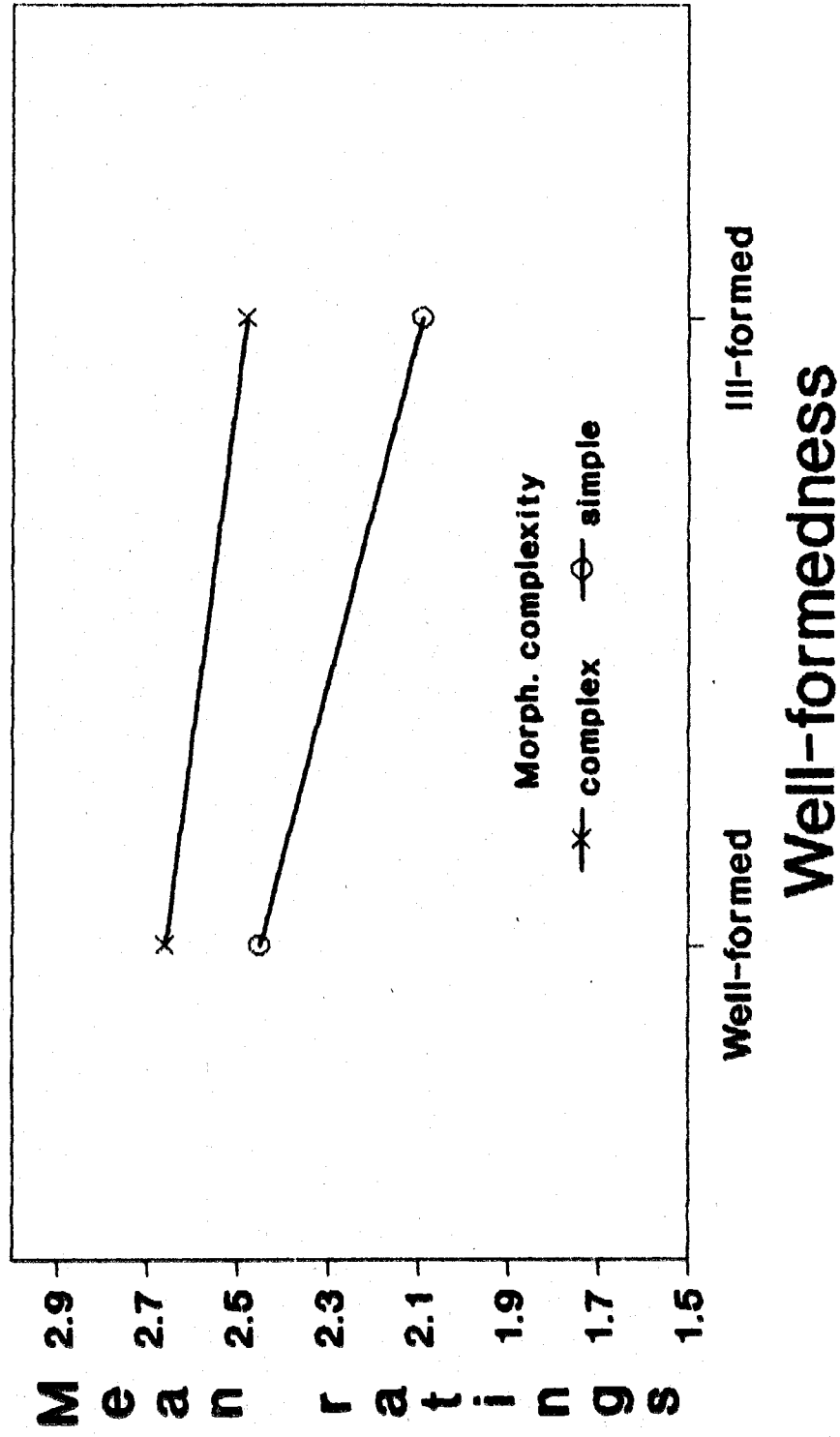
Well-formedness by Part of Speech Interaction



Ratings: 4=very likely 1=very unlikely

Figure 2

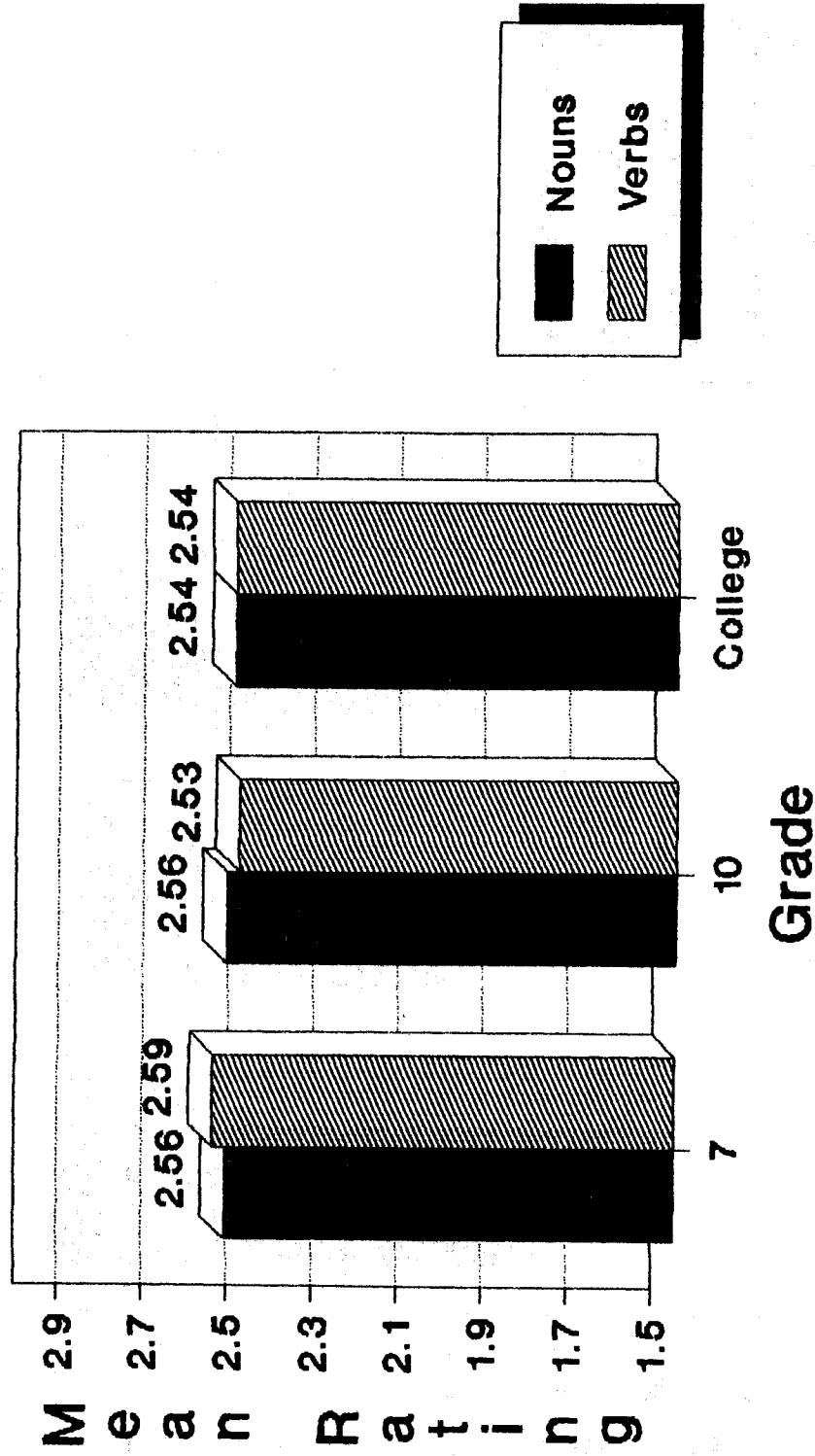
Interaction of morphological complexity and well-formedness



Ratings: 4=very likely 1=very unlikely

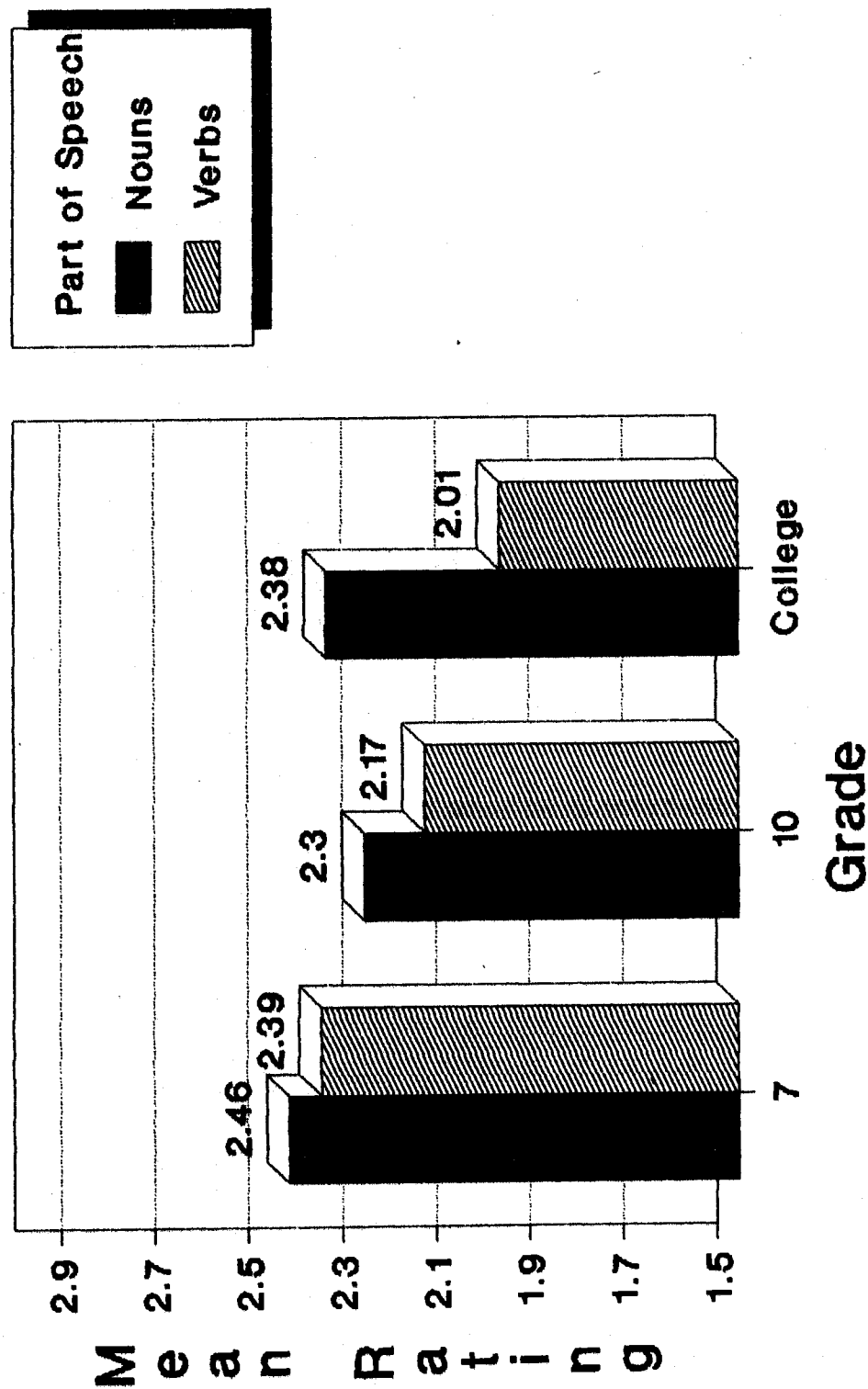
Figure 3

Grade by Well-formedness by Part of Speech



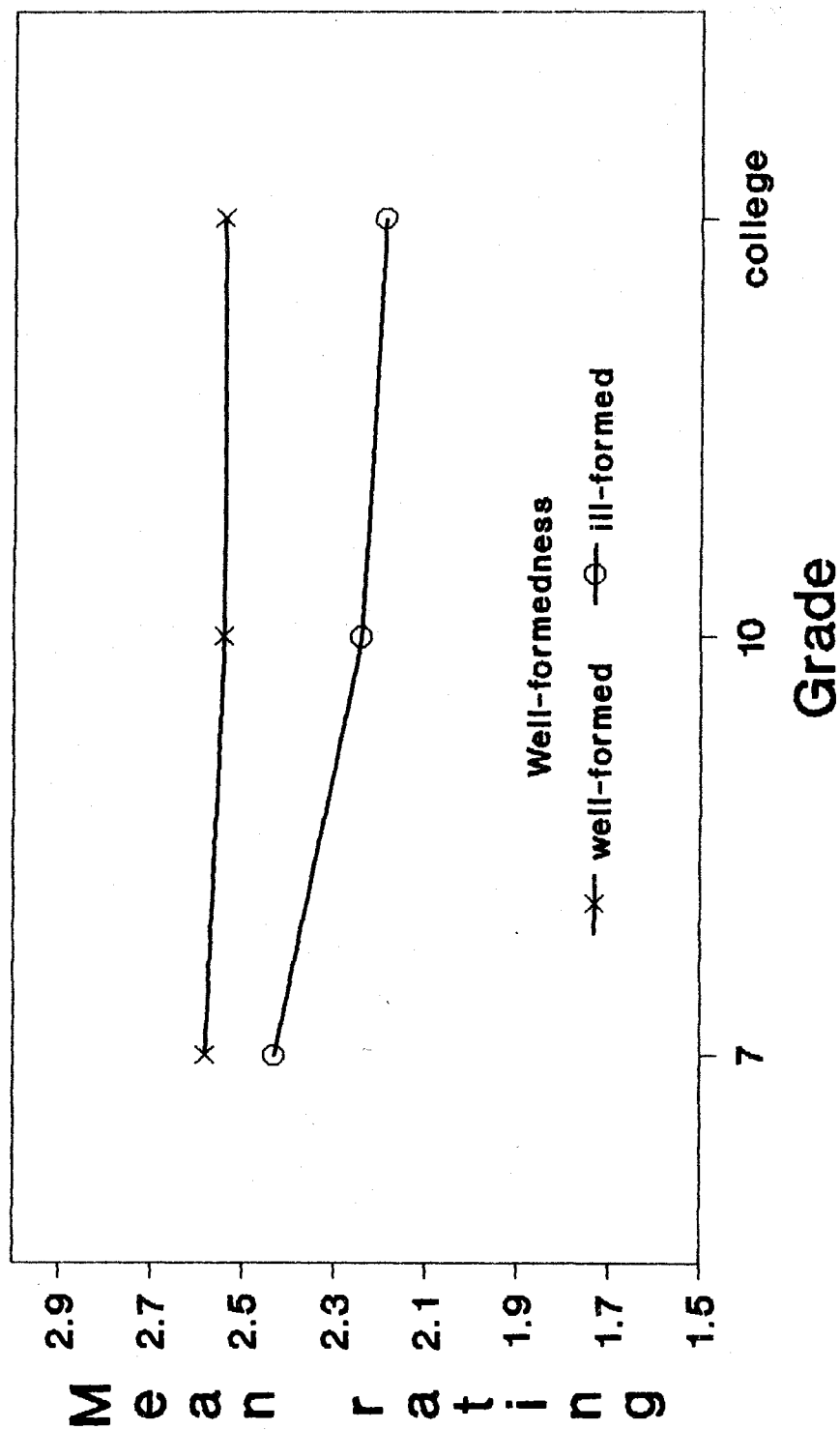
(a) Well-formed Meanings

Ratings: 4=very likely 1=very unlikely



(b) Ill-formed Meanings

Grade by well-formedness Interaction



Ratings: 4=very likely 1=very unlikely

Figure 5

This page is intentionally blank.

